



TITLE:

Levenberg-Marquardt法の局所収束性について (最適化の数理科学)

AUTHOR(S):

山下, 信雄; 福島, 雅夫

CITATION:

山下, 信雄 ...[et al]. Levenberg-Marquardt法の局所収束性について (最適化の数理科学). 数理解析研究所講究録 2000, 1174: 161-168

ISSUE DATE:

2000-10

URL:

<http://hdl.handle.net/2433/64462>

RIGHT:

Levenberg-Marquardt 法の局所収束性について

京都大学大学院情報学研究科 山下 信雄
福島 雅夫

摘要

非線形方程式 $F(x) = 0$ の解法として Levenberg-Marquardt 法 (LM 法) は非常に有効な手法である [1, 2]. 非線形方程式の解 x において F のヤコビアンが正則であるとき, 初期点を x の十分近くにとれば, LM 法で生成される点列は解に 2 次収束することが知られている. 本論文では, 解 x において F のヤコビアンが正則でない場合でも, $\|F(x)\|$ が非線形方程式に対するエラーバウンドを与えていれば, LM 法で生成される点列と解集合との距離が 0 に 2 次収束することを示す.

1 はじめに

非線形方程式

$$F(x) = 0 \quad (1)$$

の解法として, Levenberg-Marquardt 法 (以下, LM 法) は非常に有効な手法である. ここで, F は R^n から R^n への連続微分可能な関数である.

LM 法はニュートン法的一种であり, 点 x^k が与えられたとき, 探索方向として次の線形方程式の解 d^k を用いる.

$$(F'(x^k)^T F'(x^k) + \mu_k I) d = -F'(x^k)^T F(x^k) \quad (2)$$

ここで μ_k は正の定数である. 線形方程式 (2) において, $(F'(x^k)^T F'(x^k) + \mu_k I)$ は正定値行列となるため, 探索方向 d^k は必ず求めることができる. さらに, 探索方向 d^k は, メリット関数

$$\phi(x) := \frac{1}{2} \|F(x)\|^2$$

の降下方向となる. そのため, アルミホタイプのステップサイズルールと組み合わせることによって, LM 法で生成される点列の集積点は ϕ の停留点となる. また, 停留点 x^* において, $F'(x^*)$ が正則であれば,

$$0 = \nabla \phi(x^*) = F'(x^*)^T F(x^*)$$

であることより, x^* は問題 (1) の解となる. よって, このような条件が集積点において成り立つとき, LM 法は大域的収束する. (実際にはもう少し弱い条件で, ϕ の停留点が (1) の解になることを示せる.) また, ある集積点 x^* における $F'(x^*)$ の正則性と適当な μ_k の更新規則によって, LM 法で生成される点列は x^* に超一次収束する. このように, LM 法は現在知られている非線形方程式の解法の中では大域的にも局所的にもよい収束性を持つことが示されている.

最近, 集積点 x において $F'(x)$ の正則性が成り立たなくても, 超一次収束する手法が提案されている [4, 5]. そのような手法では, アルゴリズムが超一次収束するためには, $F'(x)$ の正則性の代わりに, F が局所的にエラーバウンドとなることを必要としている. ここで, 局所エラーバウンド性は以下のように定義されている性質である.

Definition 1.1 集合 X を問題 (1) の解集合とし, $N \subseteq R^n$ かつ $X \cap N \neq \emptyset$ とする. このとき次の不等式を満たす正の定数 c が存在するとき, F は集合 N 上で局所エラーバウンドになるという.

$$c \operatorname{dist}(x, X) \leq \|F(x)\| \quad \forall x \in N$$

ある解 x^* において $F'(x)$ が正則であれば, x^* は (1) の孤立して解であり, F は点 x^* の適当な近傍上で局所的エラーバウンドとなる. よって F が局所的エラーバウンドとなるという条件は, 集積点 x における $F'(x)$ の正則性よりも緩い条件である.

本論文では, F がある解 x^* の近傍上で局所的エラーバウンドになり, x^* の十分近くに初期点を選べば, LM 法によって生成される点列が解に超一次収束することを示す. さらに, この結果を線形相補性問題の再定式化によって得られる方程式系に応用することによって, 線形相補性問題に対してこれまでに提案されている手法よりも緩い条件で超一次収束することを示す.

2 LM 法の性質

この節では, 直線探索を用いない LM 法の局所的収束性を議論する. つまり, 次の反復点 x^{k+1} は

$$x^{k+1} := x^k + d^k$$

とする. ここで, LM 法の探索方向 d^k は, 線形方程式

$$(F'(x^k)^T F'(x^k) + \mu_k I) d = -F'(x^k)^T F(x^k) \quad (3)$$

の解である.

これまでの LM 法の収束率の解析には, 解の十分そばではニュートン方程式と (3) が等価になることを用いていた. ここでは, この方程式 (3) と等価な制約なし最小化問題を与え, その最小化問題の性質を用いることによって, 収束率の解析を行う.

そのために, まず, 関数 θ^k を次のように定義する.

$$\theta^k(d) = \|F'(x^k)d + F(x^k)\|^2 + \mu_k \|d\|^2$$

容易にわかるように関数 θ^k は狭義凸 2 次関数である. さらにこの関数の制約無し最小化問題

$$\min \theta^k(d) \quad (4)$$

は, θ が凸関数であることと最適性の 1 次の条件より, 方程式 (3) と等価であることがわかる. この最小化問題 (4) の性質を用いて, LM 法の局所的な性質の解析を行う.

まず, これからの解析に必要ないくつかの仮定をおく.

Assumption 2.1 問題 (1) の解集合 X は空でなく, ある $x^* \in X$ において次の条件 (a), (b) が成り立つ.

(a) 以下の不等式が成り立つような定数 $b \in (0, \infty)$, $c_1 \in (0, \infty)$ が存在する.

$$\|F'(y)(x - y) - (F(x) - F(y))\| \leq c_1 \|x - y\|^2 \quad \forall x, y \in N(x^*, b) := \{x \mid \|x - x^*\| \leq b\}$$

(b) $\|F(x)\|$ は $N(x^*, b)$ 上で局所的エラーバウンドとなる. つまり, 次の不等式が成り立つ正の定数 c_2 が存在する.

$$c_2 \text{dist}(x, X) \leq \|F(x)\| \quad \forall x \in N(x^*, b)$$

仮定 (a) は F が連続微分可能で F' がリプシッツ連続であれば成り立つ。(しかしながら, (a) は一般の微分不可能な関数では成り立たないので, この仮定を外すことは今後の課題である。) また仮定 (a) より, 次の不等式を満たす正の定数 L が存在する.

$$\|F(x) - F(y)\| \leq L\|x - y\| \quad \forall x, y \in N(x^*, b) \quad (5)$$

(b) が本論文で重要な役割を果たす仮定である. 関数 F が線形関数 (または区分的線形関数) であれば (b) が成り立つことが知られている. (この性質により線形相補性問題と等価な方程式が局所的エラーバウンドを与える.)

また, アルゴリズムにおいて用いる点列 $\{\mu_k\}$ は以下のように更新されると仮定する.

Assumption 2.2 各 k に対して $\mu_k := \|F(x^k)\|^2$ である.

この仮定は μ_k が 0 に収束する速さを与えている. 実際, 式 (5) より, μ_k は点列が解に収束する 2 乗の速さで収束する. なお, 実際にアルゴリズムを実装する場合は, 適当な正の定数 α, β を用いて,

$$\mu_k = \alpha \min\{\|F(x^k)\|^2, \beta\}$$

とすることもできるが, 解析を容易にするために, ここでは仮定 2.2 のように定義する.

これらの仮定のもと次の性質を示すことができる. 以下では, 各 k に対して, \bar{x}^k は

$$\|x^k - \bar{x}^k\| = \text{dist}(x^k, X)$$

を満たす X の点を表すものとする.

Lemma 2.1 仮定 2.1 が成り立つとする. さらに, $x^k \in N(x^*, \frac{b}{2})$ かつ d^k を方程式 (2) の解とする. このとき,

$$\begin{aligned} \|d^k\| &\leq c_3 \text{dist}(x^k, X) \\ \|F'(x^k)d^k + F(x^k)\| &\leq c_4 \text{dist}(x^k, X) \end{aligned}$$

が成り立つ. ここで, $c_3 = \sqrt{\frac{c_1^2 + c_2^2}{c_2^2}}$, $c_4 = \sqrt{c_1^2 + L^2}$ である.

証明. d^k は最小化問題 (4) の解であるので,

$$\theta^k(d^k) \leq \theta^k(\bar{x}^k - x^k)$$

である. また, $x^k \in N(x^*, \frac{b}{2})$ であるので

$$\|\bar{x}^k - x^*\| \leq \|\bar{x}^k - x^k\| + \|x^* - x^k\| \leq \|x^* - x^k\| + \|x^* - x^k\| \leq b$$

となる. よって, $\bar{x}^k \in N(x^*, b)$ である. このため, θ^k の定義と仮定 2.1 (a) より,

$$\begin{aligned} \|d^k\|^2 &\leq \frac{1}{\mu_k} \theta^k(d^k) \\ &\leq \frac{1}{\mu_k} \theta^k(\bar{x}^k - x^k) \\ &= \frac{1}{\mu_k} (\|F'(x^k)(\bar{x}^k - x^k) + F(x^k)\|^2 + \mu_k \|\bar{x}^k - x^k\|^2) \\ &\leq \frac{1}{\mu_k} (c_1^2 \|x^k - \bar{x}^k\|^4 + \mu_k \|x^k - \bar{x}^k\|^2) \\ &\leq \left(\frac{c_1^2 \|x^k - \bar{x}^k\|^2}{\mu_k} + 1 \right) \|x^k - \bar{x}^k\|^2 \end{aligned}$$

となる. 一方, 仮定 2.1 (b), 仮定 2.2 より,

$$\mu_k = \|F(x^k)\|^2 \geq c_2^2 \|x^k - \bar{x}^k\|^2$$

であるので, この不等式を上記の不等式に代入して式変形すると,

$$\|d^k\| \leq \sqrt{\frac{c_1^2 + c_2^2}{c_2^2}} \|x^k - \bar{x}^k\|$$

を得る.

次に 2 番目の不等式を示す. 1 番目の不等式を示したときと同様にして,

$$\begin{aligned} \|F'(x^k)d^k + F(x^k)\|^2 &\leq \theta^k(d^k) \\ &\leq \theta^k(\bar{x}^k - x^k) \\ &\leq c_1^2 \|x^k - \bar{x}^k\|^4 + \mu_k \|x^k - \bar{x}^k\|^2 \end{aligned}$$

を得る. ここで, (5), 仮定 2.2 より,

$$\mu_k = \|F(x^k)\|^2 = \|F(x^k) - F(\bar{x}^k)\|^2 \leq L^2 \|x^k - \bar{x}^k\|^2$$

であるので,

$$\|F'(x^k)d^k + F(x^k)\| \leq \sqrt{c_1^2 + L^2} \|x^k - \bar{x}^k\|^2$$

を得る. □

次のこの性質を用いて, すべての k に対して $x^k \in N(x^*, b/2)$ であれば, $\text{dist}(x^k, X)$ が 0 に 2 次収束することを示す.

Lemma 2.2 $x^k, x^{k-1} \in N(x^*, b/2)$ であれば,

$$\text{dist}(x^k, X) \leq c_5 \text{dist}(x^{k-1}, X)^2 \tag{6}$$

が成り立つ. ここで, $c_5 = (c_1 c_3^2 + c_4)/c_2$ である.

証明. $x^k, x^{k-1} \in N(x^*, \frac{b}{2})$ かつ $x^k = x^{k-1} + d^{k-1}$ であるので, 仮定 2.1 (a), (b) と補題 2.1 より

$$\begin{aligned} c_2 \text{dist}(x^k, X) &= c_2 \text{dist}(x^{k-1} + d^{k-1}, X) \\ &\leq \|F(x^{k-1} + d^{k-1})\| \\ &\leq \|F'(x^{k-1})d^{k-1} + F(x^{k-1})\| + c_1 \|d^{k-1}\|^2 \\ &\leq c_4 \text{dist}(x^{k-1}, X)^2 + c_1 c_3^2 \text{dist}(x^{k-1}, X)^2 \\ &= (c_1 c_3^2 + c_4) \text{dist}(x^{k-1}, X)^2 \end{aligned}$$

となる. よって, $c_5 = (c_1 c_3^2 + c_4)/c_2$ とすれば, (6) を得る. □

この補題より, すべての k に対して $x^k \in N(x^*, \frac{b}{2})$ であれば, $\{\text{dist}(x^k, X)\}$ が 0 に 2 次収束することがわかる. そこで, 以下の補題において, すべてに k に対して $x^k \in N(x^*, \frac{b}{2})$ となるための十分条件を与える.

Lemma 2.3 $r := \min\{\frac{b}{2+4c_3}, \frac{1}{2c_5}\}$ とする. このとき $x^0 \in N(x^*, r)$ であれば, すべての k に対して $x^k \in N(x^*, b/2)$ となる.

証明. すべての k に対して, $x^l \in N(x^*, \frac{b}{2})$, $l = 0, 1, \dots, k$ であれば, $x^{k+1} \in N(x^*, \frac{b}{2})$ となることを示す. このことが示されれば, $x^0 \in N(x^*, r) \subseteq N(x^*, b/2)$ であるので, 題意が満たされたことになる.

ここで, $k = 0$, $k \geq 1$ の 2 つ場合に分けて示す.

$k = 0$ のとき: このとき, 補題 2.1 より

$$\begin{aligned} \|x^1 - x^*\| &\leq \|x^0 + d^0 - x^*\| \leq \|x^0 - x^*\| + \|d^0\| \leq r + c_3 \text{dist}(x^0, X) \\ &\leq r + c_3 \|x^0 - x^*\| \leq (1 + c_3)r \end{aligned} \quad (7)$$

を得る. $(1 + c_3)r \leq \frac{b}{2}$ なので, $x^1 \in N(x^*, b/2)$ となる.

$k \geq 1$ のとき: このとき $0 \leq l \leq k$ であるすべての l に対して, $x^l \in N(x^*, b/2)$ である. よって, 補題 2.2 より, $1 \leq l \leq k$ である l に対して

$$\text{dist}(x^l, X) \leq c_5 \text{dist}(x^{l-1}, X)^2 \leq \dots \leq c_5^{2^l-1} \|x^0 - x^*\|^{2^l} \leq r \left(\frac{1}{2}\right)^{2^l-1}$$

が成り立つ. ここで最後の不等式は $r \leq \frac{1}{2c_5}$ であることを用いた. さらに補題 2.1 より, $1 \leq l \leq k$ であるすべての l に対して

$$\|d^l\| \leq c_3 \text{dist}(x^l, X) \leq c_3 r \left(\frac{1}{2}\right)^{2^l-1} \leq c_3 r \left(\frac{1}{2}\right)^{2l-1} \quad (8)$$

となる. よって,

$$\begin{aligned} \|x^{k+1} - x^*\| &\leq \|x^1 - x^*\| + \sum_{l=1}^k \|d^l\| \\ &\leq (1 + c_3)r + c_3 r \sum_{l=1}^k \left(\frac{1}{2}\right)^{2l-1} \\ &\leq (1 + 2c_3)r \\ &\leq \frac{b}{2} \end{aligned}$$

となる. ここで, 2 番目の不等式は (7) を用い, 最後の不等式は $r \leq \frac{b}{2+4c_3}$ であることを用いた. よって $x^{k+1} \in N(x^*, \frac{b}{2})$ となる.

以上より, 題意は示された. \square

補題 2.2, 2.3 を用いれば, 本節の主題である次の定理が示せる.

Theorem 2.1 仮定 2.1 が成り立ち, $r := \min\{\frac{b}{2+4c_3}, \frac{1}{2c_5}\}$ とする. さらに $x^0 \in N(x^*, r)$ として LM 法で生成された点列を $\{x^k\}$ とする. このとき $\{\text{dist}(x^k, X)\}$ は 0 に 2 次収束する. ここで X は問題 (1) の解集合である. さらに, 点列 $\{x^k\}$ はある解 $\hat{x} \in N(x^*, b/2)$ に収束する.

証明. 定理の前半は補題 2.2, 2.3 より明らかである.

点列 $\{x^k\}$ がある解 $\hat{x} \in N(x^*, b/2)$ に収束することを示す. このことを示すには, $\{\text{dist}(x^k, X)\}$ が 0 に収束することと $\{x^k\} \subset N(x^*, b/2)$ より, $\{x^k\}$ がある点に収束することを示せば十分である. (8) より, すべての $k \geq 1$ に対して,

$$\|d^k\| \leq c_3 r \left(\frac{1}{2}\right)^{2k-1}$$

が成り立つ。このとき、 $p \geq q$ である正整数 p, q に対して

$$\|x^p - x^q\| \leq \sum_{i=q}^{p-1} \|d^i\| \leq \sum_{i=q}^{\infty} \|d^i\| \leq c_3 r \sum_{i=q}^{\infty} \left(\frac{1}{2}\right)^{2i-1} \leq 3c_3 r \left(\frac{1}{2}\right)^{2q}$$

となる。よって、 $\{x^k\}$ はコーシー列になるので、ある点に収束する。 \square

3 大域的収束性

前節では LM 法の収束率の解析をおこなった。ここでは、LM 法にアルミホのステップサイズルールを組み合わせたアルゴリズムを提案し、そのアルゴリズムの大域的収束性を議論する。なおステップサイズをきめるために用いるメリット関数は、1 節で定義した

$$\phi(x) = \frac{1}{2} \|F(x)\|^2$$

を用いる。

ここでは以下のアルゴリズムを提案する。

LM 法

ステップ 0 : 適当にパラメータ $\alpha, \beta, \gamma \in (0, 1)$ を選ぶ。初期点 x^0 を適当に選び、 $\mu_0 = \|F(x^0)\|^2$ とする。 $k = 0$ とする。

ステップ 1 : x^k が終了条件をみたせば終了する。

ステップ 2 : 次の線形方程式の解 d^k を求める。

$$(F'(x^k)^T F'(x^k) + \mu_k I) d = -F'(x^k)^T F(x^k)$$

もし

$$\|F(x^k + d^k)\| \leq \gamma \|F(x^k)\|$$

ならば、 $x^{k+1} = x^k + d^k$ としステップ 4 へ。さもなければステップ 3 へ。

ステップ 3 : 次の不等式を満たす最小の非負の整数 m を求め、 $x^{k+1} = x^k + \beta^m d^k$ とする。

$$\phi(x^k + \beta^m d^k) - \phi(x^k) \leq \alpha \beta^m \nabla \phi(x^k)^T d^k$$

ステップ 4 : $\mu_{k+1} = \|F(x^{k+1})\|^2$ とする。 $k = k + 1$ として、ステップ 1 へ。

このアルゴリズムに対して次の定理がなりたつ。

Theorem 3.2 $\{x^k\}$ を LM 法で生成された点列とする。このとき $\{x^k\}$ の任意の集積点は ϕ の停留点となる。さらに $\{x^k\}$ の集積点が問題 (1) の解を少なくとも 1 つ含むとする。その解 x^* に対して仮定 2.1 が成り立てば、点列 $\{\text{dist}(x^k, X)\}$ は 0 に 2 次収束する。

証明. $\nabla \phi(x^k) \neq 0$ であれば $d^k \neq 0$ であり、

$$\nabla \phi(x^k)^T d^k = (F'(x^k)^T F(x^k))^T d^k = -((F'(x^k)^T F'(x^k) + \mu_k) d^k)^T d^k < 0$$

である。よって、ステップ3のルールより、 $\{\phi(x^k)\}$ は単調に減少する。つまり μ_k は単調に減少する点列である。このとき、 $\mu_k \rightarrow 0$ であれば、 $F(x^k) \rightarrow 0$ であるので、任意の集積点は問題(1)の解となる。また、その集積点は ϕ の停留点となる。一方、 $\liminf \mu_k > 0$ の場合も、標準的な証明方法を用いれば、 $\{x^k\}$ の任意の集積点 x^* が ϕ の停留点となることが示せる。

さらに $\{x^k\}$ の集積点 x^* が問題(1)の解であれば、

$$\|x^{\bar{k}} - x^*\| \leq r$$

かつ、

$$F(x^{\bar{k}}) \leq \frac{c_2^2 \gamma}{c_5 L} \quad (9)$$

となる \bar{k} が存在する。ここで、 r は定理2.1によって与えられた定数である。このとき、 $y^0 = x^{\bar{k}}$ として直線探索を用いないLM法で生成された点列を $\{y^l\}$ とする。定理2.1より、 $\text{dist}(y^l, X)$ は0に2次収束する。そのため、この定理を示すには $x^{\bar{k}+l} = y^l$ となること、つまり、すべての l に対して、

$$\|F(y^{l+1})\| \leq \gamma \|F(y^l)\|$$

となることを示せば十分である。

ここで、補題2.2, 2.3より、すべての l に対して、

$$\text{dist}(y^{l+1}, X) \leq c_5 \text{dist}(y^l, X)^2$$

が成り立つことに注意する。このことと(5)、仮定2.1(b)より、

$$\|F(y^{l+1})\| = \|F(y^{l+1}) - F(\bar{y}^{l+1})\| \leq L \text{dist}(y^{l+1}, X) \leq c_5 L \text{dist}(y^l, X)^2 \leq \frac{c_5 L \|F(y^l)\|}{c_2^2} \|F(y^l)\| \quad (10)$$

を得る。(9)より、

$$\frac{c_5 L \|F(y^0)\|}{c_2^2} \leq \gamma$$

である。式(10)と $\gamma < 1$ であることより、すべての l に対して、

$$\frac{c_5 L \|F(y^0)\|}{c_2^2} \leq \gamma$$

であり、

$$\|F(y^{l+1})\| \leq \gamma \|F(y^l)\|$$

となることが示せる。よって、題意は満たされた。 \square

4 線形相補性問題への応用

この節では、LM法の線形相補性問題(LCP(M, q))への応用を考える。LCP(M, q)とは、与えられた $n \times n$ 行列 M と n 次元ベクトル q に対して、

$$x \geq 0, \quad Mx + q \geq 0, \quad x^T(Mx + q) = 0$$

となる点 $x \in R^n$ を求める問題である. これまでに, $LCP(M, q)$ と等価な方程式系はいくつか提案されている. ここでは Fischer-Brumeister 関数を元にした次の方程式を考える [3].

$$H(x) = 0$$

このとき, H は $\{x \mid \|H(x)\| \leq \varepsilon\}$ 上で局所的エラーバウンドを与えることが知られている [3]. ここで ε は十分小さい正の定数である. さらに, 点 x が非退化点であれば, H は 2 回微分可能である. よって, この方程式系に対して LM 法を適用したとき, 以下の性質が成り立つ.

Theorem 4.3 行列 M が P_0 行列であるとする. このとき LM 法で生成された点列の任意の集積点は $LCP(M, q)$ の解である. さらに, 集積点の 1 つが非退化解であるとする. このとき点列は $LCP(M, q)$ の解集合に 2 次収束する. \square

これまでに $LCP(M, q)$ に対しては,

- (a) 行列 M が半正定値で, 集積点の 1 つが非退化解である.
- (b) 行列 M が P_0 行列で, 集積点の 1 つである種の正則性が成り立っている.

のうち 1 つを満たしたときに, 超一次収束するアルゴリズムが提案されている [4, 5, 6]. 本論文では, (a) の場合に対して, 扱う行列が P_0 行列であっても超一次収束することを示している. (本文中では示していないが, (b) の場合でも超一次収束することは示せる.)

参考文献

- [1] D. P. Bertsekas, *Nonlinear Programming*, Athena Scientific, Massachusetts, 1995.
- [2] F. Facchinei and C. Kanzow, A nonsmooth inexact Newton Method for the solution of large-scale nonlinear complementarity problems, *Mathematical Programming* Vol. 76, pp. 493-512, 1997.
- [3] A. Fischer, An NCP-function and its use for the solution of complementarity problems, *Recent Advances in Nonsmooth Optimization*, D. Z. Du, L. Qi and R. S. Womersley (eds.) World Scientific, Singapore, pp. 88-105, 1995.
- [4] P. Tseng, Error bounds and superlinear convergence analysis of some Newton-type methods in optimization, to appear in *Applications and Algorithms of Complementarity*, M.C. Ferris, O.L. Mangasarian and J.-S. Pang (eds.), Kluwer Academic Publishers.
- [5] N. Yamashita and M. Fukushima, The proximal point algorithm with genuine superlinear convergence for the monotone complementarity problem, to appear in *SIAM Journal on Optimization*.
- [6] N. Yamashita, J. Imai and M. Fukushima, The proximal point algorithm for the P_0 complementarity problem, to appear in *Applications and Algorithms of Complementarity*, M.C. Ferris, O.L. Mangasarian and J.-S. Pang (eds.), Kluwer Academic Publishers.